

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT: BRASH, ALAN R.  
BOEGLIN, WILLIAM E.  
JISAKA, MITSUO

10 (ii) TITLE OF INVENTION: LIPOXYGENASE PROTEINS AND  
NUCLEIC ACIDS

10 (iii) NUMBER OF SEQUENCES: 36

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: ARLES A. TAYLOR, JR.

(B) STREET: SUITE 1400, UNIVERSITY TOWER, 3100 TOWER  
BOULEVARD

15 (C) CITY: DURHAM

(D) STATE: NORTH CAROLINA

(E) COUNTRY: USA

(F) ZIP: 27707

20 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB storage

(B) COMPUTER: IBM PC/XT/AT compatible

(C) OPERATING SYSTEM: Windows 3.1

(D) SOFTWARE: WORD PERFECT 6.1 and ASCII

25 (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/061,768

(B) FILING DATE: April 16, 1998

(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA: None

(A) APPLICATION NUMBER:

30 (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: ARLES A. TAYLOR, JR.

(B) REGISTRATION NUMBER: 39,395

(C) REFERENCE/DOCKET NUMBER: 1242/5

35 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (919) 493-8000

(B) TELEFAX: (919) 419-0383

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2685 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGGGCGTGTGTC CCAGGGGGAG CCCCCGCTCTG CAGCCCTGTG CGCCCGTAGAG AGCTGGACTT 60

5 AGGCTGGCAG C ATG GCC GAG TTC AGG GTC AGG GTG TCC ACC GGA GAA GCC 110  
Met Ala Glu Phe Arg Val Arg Val Ser Thr Gly Glu Ala  
1 5 10

TTC GGG GCT GGC ACA TGG GAC AAA GTG TCT GTC AGC ATC GTG GGG ACC 158  
10 Phe Gly Ala Gly Thr Trp Asp Lys Val Ser Val Ser Ile Val Gly Thr  
15 20 25

CGG GGA GAG AGC CCC CCA CTG CCC CTG GAC AAT CTC GGC AAG GAG TTC 206  
Arg Gly Glu Ser Pro Pro Leu Pro Leu Asp Asn Leu Gly Lys Glu Phe  
15 30 35 40 45

ACT GCG GGC GCT GAG GAG GAC TTC CAG GTG ACG CTC CCG GAG GAC GTA 254  
Thr Ala Gly Ala Glu Glu Asp Phe Gln Val Thr Leu Pro Glu Asp Val  
50 55 60

20 GGC CGA GTG CTG CTG CGC GTG CAC AAG GCG CCC CCA GTG CTG CCC 302  
Gly Arg Val Leu Leu Leu Arg Val His Lys Ala Pro Pro Val Leu Pro  
65 70 75

25 CTG CTG GGG CCC CTG GCC CCG GAT GCC TGG TTC TGC CGC TGG TTC CAG 350  
Leu Leu Gly Pro Leu Ala Pro Asp Ala Trp Phe Cys Arg Trp Phe Gln  
80 85 90

30 CTG ACA CCG CCG CGG GGC CAC CTC CTC TTC CCC TGC TAC CAG TGG 398  
Leu Thr Pro Pro Arg Gly Gly His Leu Leu Phe Pro Cys Tyr Gln Trp  
95 100 105

35 CTG GAG GGG GCG GGG ACC CTG GTG CTG CAG GAG GGT ACA GCC AAG GTG 446  
Leu Glu Gly Ala Gly Thr Leu Val Leu Gln Glu Gly Thr Ala Lys Val  
110 115 120 125

TCC TGG GCA GAC CAC CAC CCT GTG CTC CAG CAA CAG CGC CAG GAG GAG 494  
Ser Trp Ala Asp His His Pro Val Leu Gln Gln Gln Arg Gln Glu Glu

130 135 140

CTT CAG GCC CGG CAG GAG ATG TAC CAG TGG AAG GCT TAC AAC CCA GGT 542  
Leu Gln Ala Arg Gln Glu Met Tyr Gln Trp Lys Ala Tyr Asn Pro Gly  
145 150 155

5 TGG CCT CAC TGC CTG GAT GAA AAG ACA GTG GAA GAC TTG GAG CTC AAT 590  
Trp Pro His Cys Leu Asp Glu Lys Thr Val Glu Asp Leu Glu Leu Asn  
160 165 170

10 ATC AAA TAC TCC ACA GCC AAG AAT GCC AAC TTT TAT CTA CAA GCT GGC 638  
Ile Lys Tyr Ser Thr Ala Lys Asn Ala Asn Phe Tyr Leu Gln Ala Gly  
175 180 185

TCT GCT TTT GCA GAG ATG AAA ATC AAG GGG TTG CTG GAC CGC AAG GGG 686  
15 Ser Ala Phe Ala Glu Met Lys Ile Lys Gly Leu Leu Asp Arg Lys Gly  
190 195 200 205

CTC TGG AGG AGT CTG AAT GAG ATG AAA AGG ATC TTC AAC TTC CGG AGG 734  
Leu Trp Arg Ser Leu Asn Glu Met Lys Arg Ile Phe Asn Phe Arg Arg  
20 210 215 220

ACC CCA GCA GCT GAG CAC GCA TTT GAG CAC TGG CAG GAG GAT GCC TTC 782  
Thr Pro Ala Ala Glu His Ala Phe Glu His Trp Gln Glu Asp Ala Phe  
225 230 235

25 TTC GCC TCC CAG TTC CTG AAT GGT CTC AAC CCT GTC CTG ATC CGC CGC 830  
Phe Ala Ser Gln Phe Leu Asn Gly Leu Asn Pro Val Leu Ile Arg Arg  
240 245 250

30 TGT CAC TAC CTC CCA AAG AAC TTC CCC GTC ACT GAT GCC ATG GTG GCC 878  
Cys His Tyr Leu Pro Lys Asn Phe Pro Val Thr Asp Ala Met Val Ala  
255 260 265

TCA TTG TTG GGT CCT GGG ACC AGC TTG CAG GCT GAG CTA GAG AAG GGC 926  
35 Ser Leu Leu Gly Pro Gly Thr Ser Leu Gln Ala Glu Leu Glu Lys Gly  
270 275 280 285

TCC CTG TTC TTG GTG GAT CAC GGC ATC CTC TCT GGC ATC CAG ACC AAT 974  
Ser Leu Phe Leu Val Asp His Gly Ile Leu Ser Gly Ile Gln Thr Asn  
40 290 295 300

GTC ATT AAT GGG AAG CCG CAG TTC TCT GCG GCC CCA ATG ACC CTG CTA 1024  
Val Ile Asn Gly Lys Pro Gln Phe Ser Ala Ala Pro Met Thr Leu Leu  
305 310 315

5 TAC CAG AGC CCA GGC TGC GGG CCG CTG CTG CCT CTC GCC ATC CAG CTC 1070  
Tyr Gln Ser Pro Gly Cys Gly Pro Leu Leu Pro Leu Ala Ile Gln Leu  
320 325 330

AGC CAG ACC CCC GGC CCA AAC AGC CCC ATC TTC CTG CCC ACT GAT GAC 1118  
10 Ser Gln Thr Pro Gly Pro Asn Ser Pro Ile Phe Leu Pro Thr Asp Asp  
335 340 345

AAG TGG GAC TGG TTG CTG GCC AAG ACC TGG GTG CGC AAT GCC GAG TTC 1166  
Lys Trp Asp Trp Leu Leu Ala Lys Thr Trp Val Arg Asn Ala Glu Phe  
15 350 355 360 365

TCC TTC CAT GAG GCC CTC ACG CAC CTG CTG CAC TCA CAT CTG CTG CCT 1214  
Ser Phe His Glu Ala Leu Thr His Leu Leu His Ser His Leu Leu Pro  
370 375 380

20 GAG GTC TTC ACC CTG GCT ACC CTG CGT CAG CTG CCC CAC TGC CAC CCT 1262  
Glu Val Phe Thr Leu Ala Thr Leu Arg Gln Leu Pro His Cys His Pro  
385 390 395

25 CTC TTC AAG CTG CTG ATC CCG CAC ACC CGA TAC ACC CTG CAC ATC AAC 1310  
Leu Phe Lys Leu Leu Ile Pro His Thr Arg Tyr Thr Leu His Ile Asn  
400 405 410

ACA CTC GCC CGG GAG CTG CTT ATC GTG CCA GGG CAG GTG GTG GAC AGG 1358  
30 Thr Leu Ala Arg Glu Leu Leu Ile Val Pro Gly Gln Val Val Asp Arg  
415 420 425

TCC ACA GGC ATC GGC ATT GAA GGC TTC TCT GAG TTG ATA CAG AGG AAC 1406  
Ser Thr Gly Ile Gly Ile Glu Gly Phe Ser Glu Leu Ile Gln Arg Asn  
35 430 435 440 445

ATG AAG CAG CTG AAC TAT TCT CTC CTG TGT CTG CCT GAG GAT ATC CGG 1454  
Met Lys Gln Leu Asn Tyr Ser Leu Leu Cys Leu Pro Glu Asp Ile Arg  
450 455 460

40

ACC CGA GGA GTT GAA GAC ATC CCA GGC TAC TAC TAC CGT GAT GAT GGG 1502

Thr Arg Gly Val Glu Asp Ile Pro Gly Tyr Tyr Arg Asp Asp Gly

465 470 475

5 ATG CAG ATT TGG GGT GCA GTG GAA CGC TTT GTC TCT GAA ATC ATC GGT 1550

Met Gln Ile Trp Gly Ala Val Glu Arg Phe Val Ser Glu Ile Ile Gly

480 485 490

ATC TAC TAC CCA AGT GAT GAG TCT GTC CAA GAT GAC AGA GAG CTC CAG 1598

10 Ile Tyr Tyr Pro Ser Asp Glu Ser Val Gln Asp Asp Arg Glu Leu Gln

495 500 505

GCC TGG GTC AGA GAG ATC TTC TCC AAG GGC TTC CTA AAC CAG GAG AGC 1646

Ala Trp Val Arg Glu Ile Phe Ser Lys Gly Phe Leu Asn Gln Glu Ser

15 510 515 520 525

TCA GGT ATC CCT TCC TCA CTG GAG ACC CGG GAA GCC CTG GTG CAG TAT 1694

Ser-Gly-Ile Pro Ser-Ser Leu Glu-Thr-Arg-Glu-Ala Leu Val Gln-Tyr

530 535 540

20 GTC ACC ATG GTG ATA TTC ACC TGC TCA GCC AAG CAT GCG GCT GTC AGT 1742

Val Thr Met Val Ile Phe Thr Cys Ser Ala Lys His Ala Ala Val Ser

545 550 555

25 GCA GGG CAG TTT GAC TCC TGT GCT TGG ATG CCC AAC CTG CCA CCC AGC 1790

Ala Gly Gln Phe Asp Ser Cys Ala Trp Met Pro Asn Leu Pro Pro Ser

560 565 570

ATG CAG CTG CCA CCA CCC ACC TCC AAA GGC CTG GCA ACA TGC GAG GGC 1838

30 Met Gln Leu Pro Pro Pro Thr Ser Lys Gly Leu Ala Thr Cys Glu Gly

575 580 585

TTC ATA GCC ACC CTC CCA CCT GTC AAT GCC ACA TGT GAT GTC ATC CTT 1886

Phe Ile Ala Thr Leu Pro Pro Val Asn Ala Thr Cys Asp Val Ile Leu

35 590 595 600 605

GCT CTC TGG TTG CTG AGC AAG GAG CCT GGA GAC CAA AGG CCC CTG GGC 1934

Ala Leu Trp Leu Leu Ser Lys Glu Pro Gly Asp Gln Arg Pro Leu Gly

610 615 620

40

ACC TAT CCG GAT GAG LAC TTC ACA GAG GAG GCC CCT CGG CGG AGC ATC 102  
Thr Tyr Pro Asp Glu His Phe Thr Glu Glu Ala Pro Arg Arg Ser Ile  
625 630 635

5 GCC ACC TTC CAG AGC CGC CTG GCC CAG ATC TCG AGG GGC ATC CAG GAG 2030  
Ala Thr Phe Gln Ser Arg Leu Ala Gln Ile Ser Arg Gly Ile Gln Glu  
640 645 650

CGG AAC CGG GGC CTG GTG CTG CCC TAC ACC TAC CTA GAC CCT CCC CTC 2078  
10 Arg Asn Arg Gly Leu Val Leu Pro Tyr Thr Tyr Leu Asp Pro Pro Leu  
655 660 665

ATC GAG AAC AGC GTC TCC ATC TAAATCCAG GGGAACACAG GCCCAGATGA 2129  
Ile Glu Asn Ser Val Ser Ile  
15 670 675

CATCCCTTG ACCACATCGC TCTAGGATAA CTGGCACCCA GAGAAAAGGA CTCCCTAGAA 2189  
20 AAAACAGGCC CCCATGTGCC TCTCCTGGGA CAACCAAGCT CTGTAACCTCA CCCCCACAC 2249

CATACACACA CACAAAAACA GAAACAAAAT CAAACAGAG AAAGCAGAAA ATCTACCAAG 2309

AACAGAGTCT CAGGACAGAA CCACTGAGTC TTTGGAGGC TCCAAGCCTC AAAGTCCCCG 2369

25 CAGAGCCCAC CTTGAGGGTT TTGCTAGTTG GTTTGTTTT GCGTTTACAG CCGTGGGGGG 2429

AAGCACATAA TCCCCCCCCA GGGCCCCACTA GCATCCACTG ATTGGACCTT ATGGTCACCC 2489

AACTCAAGGA CAGCCACCAA GAAGTGGCTG CCAAAGAGAC TGGGCGCAGT GGCTCATGCC 2549  
30 CATAATCCCA GCACTTTGGG AGATGGAGGC GGGAAAATCA TTTGAGGTCA GAAGTTCAAG 2609

GCCAGCCTGG ACGACATAGC GAGACTCCAC CTCTACCAAA AAATAAAAAT TAAAAAACAA 2669

35 AAAAAAAA AAAAAA 2685

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 676 amino acids

40 (B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Glu Phe Arg Val Arg Val Ser Thr Gly Glu Ala Phe Gly Ala

1 5 10 15

5 Gly Thr Trp Asp Lys Val Ser Val Ser Ile Val Gly Thr Arg Gly Glu  
20 25 30

Ser Pro Pro Leu Pro Leu Asp Asn Leu Gly Lys Glu Phe Thr Ala Gly  
35 40 45

10 Ala Glu Glu Asp Phe Gln Val Thr Leu Pro Glu Asp Val Gly Arg Val  
50 55 60

Leu Leu Leu Arg Val His Lys Ala Pro Pro Val Leu Pro Leu Leu Gly  
15 65 70 75 80

Pro Leu Ala Pro Asp Ala Trp Phe Cys Arg Trp Phe Gln Leu Thr Pro  
85 90 95

20 Pro Arg Gly Gly His Leu Leu Phe Pro Cys Tyr Gln Trp Leu Glu Gly  
100 105 110

Ala Gly Thr Leu Val Leu Gln Glu Gly Thr Ala Lys Val Ser Trp Ala  
115 120 125

25 Asp His His Pro Val Leu Gln Gln Arg Gln Glu Glu Leu Gln Ala  
130 135 140

30 Arg Gln Glu Met Tyr Gln Trp Lys Ala Tyr Asn Pro Gly Trp Pro His  
145 150 155 160

Cys Leu Asp Glu Lys Thr Val Glu Asp Leu Glu Leu Asn Ile Lys Tyr  
165 170 175

35 Ser Thr Ala Lys Asn Ala Asn Phe Tyr Leu Gln Ala Gly Ser Ala Phe  
180 185 190

40 Ala Glu Met Lys Ile Lys Gly Leu Leu Asp Arg Lys Gly Leu Trp Arg  
195 200 205

Ser Leu Asn Glu Met Lys Arg Ile Phe Asn Phe Arg Arg Thr Pro Ala  
210 215 220

Ala Glu His Ala Phe Glu His Trp Gln Glu Asp Ala Phe Phe Ala Ser  
225 230 235 240

Gln Phe Leu Asn Gly Leu Asn Pro Val Leu Ile Arg Arg Cys His Tyr  
5 245 250 255

Leu Pro Lys Asn Phe Pro Val Thr Asp Ala Met Val Ala Ser Leu Leu  
260 265 270

10 Gly Pro Gly Thr Ser Leu Gln Ala Glu Leu Glu Lys Gly Ser Leu Phe  
275 280 285

Leu Val Asp His Gly Ile Leu Ser Gly Ile Gln Thr Asn Val Ile Asn  
290 295 300

15 Gly Lys Pro Gln Phe Ser Ala Ala Pro Met Thr Leu Leu Tyr Gln Ser  
305 310 315 320

Pro Gly Cys Gly Pro Leu Leu Pro Leu Ala Ile Gln Leu Ser Gln Thr  
20 325 330 335

Pro Gly Pro Asn Ser Pro Ile Phe Leu Pro Thr Asp Asp Lys Trp Asp  
340 345 350

25 Trp Leu Leu Ala Lys Thr Trp Val Arg Asn Ala Glu Phe Ser Phe His  
355 360 365

Glu Ala Leu Thr His Leu Leu His Ser His Leu Leu Pro Glu Val Phe  
370 375 380

30 Thr Leu Ala Thr Leu Arg Gln Leu Pro His Cys His Pro Leu Phe Lys  
385 390 395 400

Leu Leu Ile Pro His Thr Arg Tyr Thr Leu His Ile Asn Thr Leu Ala  
35 405 410 415

Arg Glu Leu Leu Ile Val Pro Gly Gln Val Val Asp Arg Ser Thr Gly  
420 425 430

40 Ile Gly Ile Glu Gly Phe Ser Glu Leu Ile Gln Arg Asn Met Lys Gln  
435 440 445

Leu Asn Tyr Ser Leu Leu Cys Leu Pro Glu Asp Ile Arg Thr Arg Gly  
450 455 460

Val Glu Asp Ile Pro Gly Tyr Tyr Arg Asp Asp Gly Met Gln Ile  
465 470 475 480

5 Trp Gly Ala Val Glu Arg Phe Val Ser Glu Ile Ile Gly Ile Tyr Tyr  
485 490 495

Pro Ser Asp Glu Ser Val Gln Asp Asp Arg Glu Leu Gln Ala Trp Val  
500 505 510

10 Arg Glu Ile Phe Ser Lys Gly Phe Leu Asn Gln Glu Ser Ser Gly Ile  
515 520 525

Pro Ser Ser Leu Glu Thr Arg Glu Ala Leu Val Gln Tyr Val Thr Met  
15 530 535 540

Val Ile Phe Thr Cys Ser Ala Lys His Ala Ala Val Ser Ala Gly Gln  
545 550 555 560

20 Phe Asp Ser Cys Ala Trp Met Pro Asn Leu Pro Pro Ser Met Gln Leu  
565 570 575

Pro Pro Pro Thr Ser Lys Gly Leu Ala Thr Cys Glu Gly Phe Ile Ala  
580 585 590

25 Thr Leu Pro Pro Val Asn Ala Thr Cys Asp Val Ile Leu Ala Leu Trp  
595 600 605

Leu Leu Ser Lys Glu Pro Gly Asp Gln Arg Pro Leu Gly Thr Tyr Pro  
30 610 615 620

Asp Glu His Phe Thr Glu Glu Ala Pro Arg Arg Ser Ile Ala Thr Phe  
625 630 635 640

Gln Ser Arg Leu Ala Gln Ile Ser Arg Gly Ile Gln Glu Arg Asn Arg  
35 645 650 655

Gly Leu Val Leu Pro Tyr Thr Tyr Leu Asp Pro Pro Leu Ile Glu Asn  
660 665 670

40 Ser Val Ser Ile  
675

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3205 base pairs

(B) TYPE: nucleic acid

5 (C) STRANDEDNESS: double

(D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GCG AAA TGC AGG GTG AGA GTA TCC ACG GGG GAA GCC TGT GGG GCT 48

10 Met Ala Lys Cys Arg Val Arg Val Ser Thr Gly Glu Ala Cys Gly Ala

1 5 10 15

GGC ACA TGG GAC AAA GTG TCT GTC AGC ATC GTG GGA ACC CAC GGA GAG 96

Gly Thr Trp Asp Lys Val Ser Val Ser Ile Val Gly Thr His Gly Glu

15 20 25 30

AGC CCC TTA GTA CCT CTG GAC CAT CTG GGC AAG GAG TTC AGC GCC GGT 144

Ser Pro Leu Val Pro Leu Asp His Leu Gly Lys Glu Phe Ser Ala Gly

35 40 45

20 GCT GAA GAA GAC TTC GAG GTG ACG CTT CCC CAG GAC GTA GGC ACT GTG 192

Ala Glu Glu Asp Phe Glu Val Thr Leu Pro Gln Asp Val Gly Thr Val

50 55 60

25 CTG ATG CTG CGA GTC CAC AAA GCA CCC CCG GAA GTG TCC CTC CCG CTT 240

Leu Met Leu Arg Val His Lys Ala Pro Pro Glu Val Ser Leu Pro Leu

65 70 75 80

ATG TCT TTC CGT TCT GAT GCC TGG TTC TGC CGC TGG TTC GAG CTG GAG 288

Met Ser Phe Arg Ser Asp Ala Trp Phe Cys Arg Trp Phe Glu Leu Glu

30 85 90 95

TGG CTA CCT GGG GCT GCA CTC CAC TTC CCC TGT TAT CAG TGG CTG GAA 336

Trp Leu Pro Gly Ala Ala Leu His Phe Pro Cys Tyr Gln Trp Leu Glu

100 105 110

GGG GCG GGG GAG CTG GTG CTG AGA GAG GGA GCA GCA AAG GTG TCC TGG 384

35 Gly Ala Gly Glu Leu Val Leu Arg Glu Gly Ala Ala Lys Val Ser Trp

115 120 125

CAA GAC CAT CAC CCT ACA CTG CAG GAT CAG CGC CAG AAG GAG CTT GAG 432

Gln Asp His His Pro Thr Leu Gln Asp Gln Arg Gln Lys Glu Leu Glu

40 130 135 140

TCC AGG CAG AAG ATG TAC AGC TGG AAG ACT TAC ATT GAA GGT TGG CCT 480  
Ser Arg Gln Lys Met Tyr Ser Trp Lys Thr Tyr Ile Glu Gly Trp Pro  
145 150 155 160

5 CGC TGC CTT GAC CAC GAG ACT GTG AAA GAC TTG GAC CTC AAC ATC AAG 528  
Arg Cys Leu Asp His Glu Thr Val Lys Asp Leu Asp Leu Asn Ile Lys  
165 170 175

TAC TCT GCG ATG AAG AAT GCC AAA CTC TTC TTT AAA GCC CAC TCC GCG 576  
10 Tyr Ser Ala Met Lys Asn Ala Lys Leu Phe Phe Lys Ala His Ser Ala  
180 185 190

TAT ACG GAG CTG AAA GTC AAA GGG CTC CTG GAC CGC ACA GGA CTC TGG 624  
Tyr Thr Glu Leu Lys Val Lys Gly Leu Leu Asp Arg Thr Gly Leu Trp  
15 195 200 205

AGG AGT CTG AGG GAG ATG AGA AGG CTG TTT AAC TTC CGC AAG ACT CCA 672  
Arg Ser Leu Arg Glu Met Arg Arg Leu Phe Asn Phe Arg Lys Thr Pro  
20 210 215 220

GCA GCA GAG TAT GTG TTT GCA CAC TGG CAG GAA GAT GCC TTC TTC GCC 720  
Ala Ala Glu Tyr Val Phe Ala His Trp Gln Glu Asp Ala Phe Phe Ala  
225 230 235 240

TCC CAG TTC CTA AAT GGC ATC AAC CCG GTC CTG ATT CGC CGC TGT CAC 768  
25 Ser Gln Phe Leu Asn Gly Ile Asn Pro Val Leu Ile Arg Arg Cys His  
245 250 255

AGT CTC CCA AAC AAC TTC CCG GTC ACT GAT GAA ATG GTG GCC CCA GTG 816  
30 Ser Leu Pro Asn Asn Phe Pro Val Thr Asp Glu Met Val Ala Pro Val  
260 265 270

CTG GGC CCT GGA ACC AGT CTG CAG GCT GAG TTG GAG AAG GGC TCC CTG 864  
Leu Gly Pro Gly Thr Ser Leu Gln Ala Glu Leu Glu Lys Gly Ser Leu  
35 275 280 285

TTC TTG GTG GAT CAT GGC ATT CTT TCT GGA GTC CAC ACC AAC ATC CTC 912  
Phe Leu Val Asp His Gly Ile Leu Ser Gly Val His Thr Asn Ile Leu  
290 295 300

40

AAT GGA AAG CCT CAG ... TCT GCA GCC CCG ATG ACC CTG TTA CAC CAG  
Asn Gly Lys Pro Gln Phe Ser Ala Ala Pro Met Thr Leu Leu His Gln  
305 310 315 320

5 AGC TCA GGG TCC GGA CCC CTG CTT CCC ATT GCC ATC CAG CTC AAA CAG 1008  
Ser Ser Gly Ser Gly Pro Leu Leu Pro Ile Ala Ile Gln Leu Lys Gln  
325 330 335

ACT CCC GGG CCA GAC AAC CCC ATC TTC CTG CCC AGC GAT GAC ACG TGG 1056  
10 Thr Pro Gly Pro Asp Asn Pro Ile Phe Leu Pro Ser Asp Asp Thr Trp  
340 345 350

GAC TGG TTG CTG GCC AAG ACC TGG GTT CGC AAT TCT GAG TTT TAC ATC 1104  
Asp Trp Leu Leu Ala Lys Thr Trp Val Arg Asn Ser Glu Phe Tyr Ile  
15 355 360 365

CAT GAG GCT GTC ACA CAT CTG CTG CAT GCC CAT CTG ATT CCA GAA GTC 1152  
His Glu Ala Val Thr His Leu Leu His Ala His Leu Ile Pro Glu Val  
370 375 380

TTT GCC TTG GCC ACA TTA CGT CAG CTG CCT AGG TGT CAC CCT CTC TTC 1200  
20 Phe Ala Leu Ala Thr Leu Arg Gln Leu Pro Arg Cys His Pro Leu Phe  
385 390 395 400

AAG CTA TTG ATT CCT CAC ATT CGG TAC ACA CTG CAC ATC AAC ACG CTT 1248  
Lys Leu Leu Ile Pro His Ile Arg Tyr Thr Leu His Ile Asn Thr Leu  
405 410 415

25 GCC CGG GAG CTG CTC GTT GCC CCT GGG AAG TTG ATA GAC AAG TCC ACA 1296  
Ala Arg Glu Leu Leu Val Ala Pro Gly Lys Leu Ile Asp Lys Ser Thr  
420 425 430

GGC CTT GGC ACT GGG GGA TTC TCT GAC CTG ATA AAG AGA AAC ATG GAG 1344  
30 Gly Leu Gly Thr Gly Phe Ser Asp Leu Ile Lys Arg Asn Met Glu  
435 440 445

CAG CTG AAC TAC TCT GTC CTG TGT CTC CCT GAA GAT ATC CGA GCC CGA 1392  
Gln Leu Asn Tyr Ser Val Leu Cys Leu Pro Glu Asp Ile Arg Ala Arg  
35 450 455 460

GGT GTG GAA GAC ATC GGC TAC TAT TAC CGA GAT GAT GGG ATG CAG  
Gly Val Glu Asp Ile Pro Gly Tyr Tyr Tyr Arg Asp Asp Gly Met Gln  
465 470 475 480

5 ATC TGG GGG GCA ATA AAG AGC TTT GTC TCT GAA ATA GTC AGC ATC TAC 1488  
Ile Trp Gly Ala Ile Lys Ser Phe Val Ser Glu Ile Val Ser Ile Tyr  
485 490 495

TAT CCA AGT GAC ACA TCC GTC CAA GAT GAC CAA GAG CTC CAG GCC TGG 1536  
10 Tyr Pro Ser Asp Thr Ser Val Gln Asp Asp Gln Glu Leu Gln Ala Trp  
500 505 510

GTG AGG GAG ATC TTC TCT GAG GGC TTC CTC GGC CGA GAA AGC TCA GGT 1584  
Val Arg Glu Ile Phe Ser Glu Gly Phe Leu Gly Arg Glu Ser Ser Gly  
15 515 520 525

ATG CCC TCC TTG TTG GAT ACC CGG GAA GCC CTG GTC CAG TAT ATC ACC 1632  
Met Pro Ser Leu Leu Asp Thr Arg Glu Ala Leu Val Gln Tyr Ile Thr  
530 535 540

20 ATG GTG ATA TTC ACC TGC TCA GCC AAG CAT GCA GCT GTC AGT TCA GGC 1680  
Met Val Ile Phe Thr Cys Ser Ala Lys His Ala Ala Val Ser Ser Gly  
545 550 555 560

25 CAG TTC GAC TCT TGT GTT TGG ATG CCC AAT CTG CCA CCT ACC ATG CAG 1728  
Gln Phe Asp Ser Cys Val Trp Met Pro Asn Leu Pro Pro Thr Met Gln  
565 570 575

CTA CCA CCA CCT ACT TCC AAA GGC CAG GCC CGG CCT GAG AGT TTC ATA 1776  
30 Leu Pro Pro Pro Thr Ser Lys Gly Gln Ala Arg Pro Glu Ser Phe Ile  
580 585 590

GCC ACG CTC CCA GCA GTT AAT TCG TCA AGT TAT CAC ATC ATT GCT CTC 1824  
Ala Thr Leu Pro Ala Val Asn Ser Ser Ser Tyr His Ile Ile Ala Leu  
35 595 600 605

TGG CTG CTA AGC GCA GAA CCT GGG GAC CAA AGG CCC CTG GGC CAC TAT 1872  
Trp Leu Leu Ser Ala Glu Pro Gly Asp Gln Arg Pro Leu Gly His Tyr  
610 615 620

40

CCA GAT GAA CAC TTC Ala GAG GAT GCC CCC CGG CGA AGC GTG GCT GCC 110  
Pro Asp Glu His Phe Thr Glu Asp Ala Pro Arg Arg Ser Val Ala Ala  
625 630 635 640

5 TTC CAG AGA AAG CTG ATC CAG ATC TCC AAG GGC ATC AGG GAG AGG AAC 1968  
Phe Gln Arg Lys Leu Ile Gln Ile Ser Lys Gly Ile Arg Glu Arg Asn  
645 650 655

CGA GGC CTG GCA CTG CCC TAC ACC TAC CTG GAT CCT CCC CTC ATT GAG 2016  
10 Arg Gly Leu Ala Leu Pro Tyr Thr Tyr Leu Asp Pro Pro Leu Ile Glu  
660 665 670

AAC AGT GTC TCC ATC TAACATCTTG GAGAAGACAG TCCTGTGTGA CATATAGAAC 2071  
Asn Ser Val Ser Ile  
15 675

TCTTGACCAT GCCTCTCCAG GCTAAGTCCC CGTATGCTTC TCCTGGACAA CCAAGCCCCA 2131  
TCTTACACAC ACACACACAC ACACACACCT AATAAAATCG AAACAGAAAA ACCTAAACTC 2191  
20 CCACAGAAGG CAAGATCTCA CACAGCAGAG AGCCATCCAA ATGTTGGAG ACCCTGAGCT 2251  
TCAGCTCTGA TTAACGGCTT TGCTGGTTG CTTTGCTTTC TATTCCATTA ACCATGGACG 2311

GTAACAGAAA GCACAGAACCC CTGGTTCACT GCACAAAGCC ACTGAGATCT CACCCCTCACC 2371  
TGACACAAAG GCAGCTATCA TACAGGCTTA TCAGGAACAC AGGAATTGT CCAATCAAAG 2431  
CCTACCCACT AGGTCCATCG TGACCTACGA CCTCACACTG GCATGCTTTA GCTTGAGAA 2491  
30 GGGATTACTG GAGTCAGGTA CGAAGAGAAG GACAGGACGA AGGCATGGCT CCATGTGGAA 2551  
GAACATATCT GCTCTTCCAG ATGACCAGGG TAGCTCACAG CCATGTGTCA TTCTAACTCC 2611

AGAGGTCTCT AGTGGCCATG AAGACTCCAG GCATTCAAGG GATATACCAAG TAGACACCAA 2671  
AATTATACTT TTTAAGAGAG AGGATGGGCT GGAGAGATGG CTCAGCGGTT AAGAGCACTG 2731  
ACTGCTCTTC CAGAGATCCT GAGTTCAATT CCCAGCAACC ACATGGTGGC TCACAACCAT 2791  
40 CTGTAATGGG ATTGATGCC CTCTCTGGC GTGCTGAAG ACAGCGACAG TGTATGCACA 2851  
TATATAAAAT AAATAAACTT TTAAAAAAACA AAACAAGAGA GAGGGACATG CTACCAATTTC 2911  
TACCTCACTT CTTCTCAAAG CCACCCCTAA AGTGAATTGT GAACCAGGTC CCCTTTGCAG 2971

AGAGTTAGAA GATATTCTCA AACCTCTAAT ACCTTCACAT CTAAAATCCA TCTTCATTCC 3031

AAAATCCAA TATTTATAT ACACCTCTCA GTTTGGTGGG TGAGGGGTTG TTTTTGTTT 3091

5

GGTTTGGTTT GGTTGGGGTT TTGTTTTGT TTTTGATTTT GTTTTCTCT GGTCAGACT 3151

CCATGGACGT TCATTAATGT CATAAATGAG TTCATTCCAA AAAAAAAA AAAA 3205

10 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 677 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

15 (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Lys Cys Arg Val Arg Val Ser Thr Gly Glu Ala Cys Gly Ala

1 5 10 15

20 Gly Thr Trp Asp Lys Val Ser Val Ser Ile Val Gly Thr His Gly Glu

20 25 30

Ser Pro Leu Val Pro Leu Asp His Leu Gly Lys Glu Phe Ser Ala Gly

35 40 45

25 Ala Glu Glu Asp Phe Glu Val Thr Leu Pro Gln Asp Val Gly Thr Val

50 55 60

Leu Met Leu Arg Val His Lys Ala Pro Pro Glu Val Ser Leu Pro Leu

30 65 70 75 80

Met Ser Phe Arg Ser Asp Ala Trp Phe Cys Arg Trp Phe Glu Leu Glu

85 90 95

35 Trp Leu Pro Gly Ala Ala Leu His Phe Pro Cys Tyr Gln Trp Leu Glu

100 105 110

Gly Ala Gly Glu Leu Val Leu Arg Glu Gly Ala Ala Lys Val Ser Trp

115 120 125

40 Gln Asp His His Pro Thr Leu Gln Asp Gln Arg Gln Lys Glu Leu Glu

130 135 140

Ser Arg Gln Lys Met 1, Ser Trp Lys Thr Tyr Ile Glu Gly Trp Pro  
145 150 155 160

Arg Cys Leu Asp His Glu Thr Val Lys Asp Leu Asp Leu Asn Ile Lys  
5 165 170 175

Tyr Ser Ala Met Lys Asn Ala Lys Leu Phe Phe Lys Ala His Ser Ala  
180 185 190

10 Tyr Thr Glu Leu Lys Val Lys Gly Leu Leu Asp Arg Thr Gly Leu Trp  
195 200 205

Arg Ser Leu Arg Glu Met Arg Arg Leu Phe Asn Phe Arg Lys Thr Pro  
210 215 220

15 Ala Ala Glu Tyr Val Phe Ala His Trp Gln Glu Asp Ala Phe Phe Ala  
225 230 235 240

Ser Gln Phe Leu Asn Gly Ile Asn Pro Val Leu Ile Arg Arg Cys His  
245 250 255

20 Ser Leu Pro Asn Asn Phe Pro Val Thr Asp Glu Met Val Ala Pro Val  
260 265 270

Leu Gly Pro Gly Thr Ser Leu Gln Ala Glu Leu Glu Lys Gly Ser Leu  
275 280 285

25 Phe Leu Val Asp His Gly Ile Leu Ser Gly Val His Thr Asn Ile Leu  
290 295 300

Asn Gly Lys Pro Gln Phe Ser Ala Ala Pro Met Thr Leu Leu His Gln  
30 305 310 315 320

Ser Ser Gly Ser Gly Pro Leu Leu Pro Ile Ala Ile Gln Leu Lys Gln  
325 330 335

35 Thr Pro Gly Pro Asp Asn Pro Ile Phe Leu Pro Ser Asp Asp Thr Trp  
340 345 350

Asp Trp Leu Leu Ala Lys Thr Trp Val Arg Asn Ser Glu Phe Tyr Ile  
355 360 365

40 His Glu Ala Val Thr His Leu Leu His Ala His Leu Ile Pro Glu Val  
370 375 380

Phe Ala Leu Ala Thr Leu Arg Gln Leu Pro Arg Cys His Pro Leu Phe  
385 390 395 400

5 Lys Leu Leu Ile Pro His Ile Arg Tyr Thr Leu His Ile Asn Thr Leu  
405 410 415

Ala Arg Glu Leu Leu Val Ala Pro Gly Lys Leu Ile Asp Lys Ser Thr  
420 425 430

10 Gly Leu Gly Thr Gly Gly Phe Ser Asp Leu Ile Lys Arg Asn Met Glu  
435 440 445

Gln Leu Asn Tyr Ser Val Leu Cys Leu Pro Glu Asp Ile Arg Ala Arg  
15 450 455 460

Gly Val Glu Asp Ile Pro Gly Tyr Tyr Tyr Arg Asp Asp Gly Met Gln  
465 470 475 480

20 Ile Trp Gly Ala Ile Lys Ser Phe Val Ser Glu Ile Val Ser Ile Tyr  
485 490 495

Tyr Pro Ser Asp Thr Ser Val Gln Asp Asp Gln Glu Leu Gln Ala Trp  
500 505 510

25 Val Arg Glu Ile Phe Ser Glu Gly Phe Leu Gly Arg Glu Ser Ser Gly  
515 520 525

Met Pro Ser Leu Leu Asp Thr Arg Glu Ala Leu Val Gln Tyr Ile Thr  
30 530 535 540

Met Val Ile Phe Thr Cys Ser Ala Lys His Ala Ala Val Ser Ser Gly  
545 550 555 560

35 Gln Phe Asp Ser Cys Val Trp Met Pro Asn Leu Pro Pro Thr Met Gln  
565 570 575

Leu Pro Pro Pro Thr Ser Lys Gly Gln Ala Arg Pro Glu Ser Phe Ile  
580 585 590

40 Ala Thr Leu Pro Ala Val Asn Ser Ser Tyr His Ile Ile Ala Leu  
595 600 605

Trp Leu Leu Ser Ala Glu Pro Gly Asp Gln Arg Pro Leu Gly His Tyr  
610 615 620

Pro Asp Glu His Phe Thr Glu Asp Ala Pro Arg Arg Ser Val Ala Ala

625 630 635 640

5 Phe Gln Arg Lys Leu Ile Gln Ile Ser Lys Gly Ile Arg Glu Arg Asn  
645 650 655

Arg Gly Leu Ala Leu Pro Tyr Thr Tyr Leu Asp Pro Pro Leu Ile Glu

10 660 665 670

Asn Ser Val Ser Ile

675

**(2) INFORMATION FOR SEQ ID NO:5:**

**(i) SEQUENCE CHARACTERISTICS:**

15 (A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

**(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:**

20 Trp Leu Leu Ala Lys  
1 5

**(2) INFORMATION FOR SEQ ID NO:6:**

**(i) SEQUENCE CHARACTERISTICS:**

25 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

**(ix) FEATURE:**

30 (A) NAME/KEY: N=i=inosine  
(B) LOCATION: 12, 15, 18  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

**(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:**

GACGTCTGGY TNYTNGCAA A 21

35 **(2) INFORMATION FOR SEQ ID NO:7:**

**(i) SEQUENCE CHARACTERISTICS:**

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (ix) FEATURE:

- (A) NAME/KEY: N=i=inosine
- (B) LOCATION: 12, 15, 18
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GACGTCTGGY TNYTNGCAA G 21

(2) INFORMATION FOR SEQ ID NO:8:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

20 Gly Gln Leu Asp Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCAAGTGAC CATTGAGT GNCC 24

30 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCAAGTGTAC CACRCRTAYT GNCC 24

(2) INFORMATION FOR SEQ ID NO:11:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ix) FEATURE:  
(A) NAME/KEY: N=i=inosine  
(B) LOCATION: 24, 25  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TAGTCGACTG GCTTYTGGCC AAANNCTGGG TSCG 34

(2) INFORMATION FOR SEQ ID NO:12:

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCGGATCCCT CCACCAAGGNY TGSAGYTC 28

25 (2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGTATCTACT ACCCAAGTGA TGAG 24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TACCCAAGTG ATGAGTCTGT C 21

(2) INFORMATION FOR SEQ ID NO:15:

10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAAGACCTCA GGCAGCAGAT GTG 23

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCATGGAAGG AGAACTCGGC AT 22

25

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACGGATCCAG CATGGCCGAG TTCAGGGTCA G 31

35

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

5 CGGAATTCA GTCATCTGGG CCTGTGTTCC 30

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGCCCTCTCGC CATCCAGCT 19

15 (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGTTCCCTG GGATTTAGAT GGA 23

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGTATCTACT ACCCAAGTGA TGAG 24

30 (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGGGATGTCA TCTGGGCCTG T 21

(2) INFORMATION FOR SEQ ID NO:23:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AACTCACCCC CACCAACATA CACA 24

(2) INFORMATION FOR SEQ ID NO:24:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTCCCCGCTC CATCTCCCAA AGT 23

20 (2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 662 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Gly Leu Tyr Arg Ile Arg Val Ser Thr Gly Ala Ser Leu Tyr Ala

1 5 10 15

30 Gly Ser Asn Asn Gln Val Gln Leu Trp Leu Val Gly Gln His Gly Glu

20 25 30

Ala Ala Leu Gly Lys Arg Leu Trp Pro Ala Arg Gly Lys Glu Thr Glu

35 40 45

35

Leu Lys Val Glu Val F. J Glu Tyr Leu Gly Pro Leu Leu Phe Val Lys  
50 55 60

Leu Arg Lys Arg His Leu Leu Lys Asp Asp Ala Trp Phe Cys Asn Trp  
5 65 70 75 80

Ile Ser Val Gln Gly Pro Gly Ala Gly Asp Glu Val Arg Phe Pro Cys  
85 90 95

10 Tyr Arg Trp Val Glu Gly Asn Gly Val Leu Ser Leu Pro Glu Gly Thr  
100 105 110

Gly Arg Thr Val Gly Glu Asp Pro Gln Gly Leu Phe Gln Lys His Arg  
115 120 125

15 Glu Glu Glu Leu Glu Glu Arg Arg Lys Leu Tyr Arg Trp Gly Asn Trp  
130 135 140

Lys Asp Gly Leu Ile Leu Asn Met Ala Gly Ala Lys Leu Tyr Asp Leu  
20 145 150 155 160

Pro Val Asp Glu Arg Phe Leu Glu Asp Lys Arg Val Asp Phe Glu Val  
165 170 175

25 Ser Leu Ala Lys Gly Leu Ala Asp Leu Ala Ile Lys Asp Ser Leu Asn  
180 185 190

Val Leu Thr Cys Trp Lys Asp Leu Asp Asp Phe Asn Arg Ile Phe Trp  
195 200 205

30 Cys Gly Gln Ser Lys Leu Ala Glu Arg Val Arg Asp Ser Trp Lys Glu  
210 215 220

Asp Ala Leu Phe Gly Tyr Gln Phe Leu Asn Gly Ala Asn Pro Val Val  
35 225 230 235 240

Leu Arg Arg Ser Ala His Leu Pro Ala Arg Leu Val Phe Pro Pro Gly  
245 250 255

40 Met Glu Glu Leu Gln Ala Gln Leu Glu Lys Glu Leu Glu Gly Thr  
260 265 270

Leu Phe Glu Ala Asp Phe Ser Leu Leu Asp Gly Ile Lys Ala Asn Val  
275 280 285

Ile Leu Cys Ser Gln Gln His Leu Ala Ala Pro Leu Val Met Leu Lys

290 295 300

Leu Gln Pro Asp Gly Lys Leu Leu Pro Met Val Ile Gln Leu Gln Leu

5 305 310 315 320

Pro Arg Thr Gly Ser Pro Pro Pro Pro Leu Phe Leu Pro Thr Asp Pro

325 330 335

Pro Met Ala Trp Leu Leu Ala Lys Cys Trp Val Arg Ser Ser Asp Phe

340 345 350

10 Gln Leu His Glu Leu Gln Ser His Leu Leu Arg Gly His Leu Met Ala

355 360 365

Glu Val Ile Val Val Ala Thr Met Arg Cys Leu Pro Ser Ile His Pro

15 370 375 380

Ile Phe Lys Leu Ile Ile Pro His Leu Arg Tyr Thr Leu Glu Ile Asn

385 390 395 400

20 Val Arg Ala Arg Thr Gly Leu Val Ser Asp Met Gly Ile Phe Asp Gln

405 410 415

Ile Met Ser Thr Gly Gly Gly His Val Gln Leu Leu Lys Gln Ala

420 425 430

25 Gly Ala Phe Leu Thr Tyr Ser Ser Phe Cys Pro Pro Asp Asp Leu Ala

435 440 445

Asp Arg Gly Leu Leu Gly Val Lys Ser Ser Phe Tyr Ala Gln Asp Ala

30 450 455 460

Leu Arg Leu Trp Glu Ile Ile Tyr Arg Tyr Val Glu Gly Ile Val Ser

465 470 475 480

35 Leu His Tyr Lys Thr Asp Val Ala Val Lys Asp Asp Pro Glu Leu Gln

485 490 495

Thr Trp Cys Arg Glu Ile Thr Glu Ile Gly Leu Gln Gly Ala Gln Asp

500 505 510

40 Arg Gly Phe Pro Val Ser Leu Gln Ala Arg Asp Gln Val Cys His Phe

515 520 525

Val Thr Met Cys Ile Phe Thr Cys Thr Gly Gln His Ala Ser Val His

530 535 540

Leu Gly Gln Leu Asp Trp Tyr Ser Trp Val Pro Asn Ala Pro Cys Thr

5 545 550 555 560

Met Arg Leu Pro Pro Pro Thr Thr Lys Asp Ala Thr Leu Glu Thr Val

565 570 575

10 Met Ala Thr Leu Pro Asn Phe His Gln Ala Ser Leu Gln Met Ser Ile

580 585 590

Thr Trp Gln Leu Gly Arg Arg Gln Pro Val Met Val Ala Val Gly Gln

595 600 605

15 His Glu Glu Glu Tyr Phe Ser Gly Pro Glu Pro Lys Ala Val Leu Lys

610 615 620

Lys Phe Arg Glu Glu Leu Ala Ala Leu Asp Lys Glu Ile Glu Ile Arg

20 625 630 635 640

Asn Ala Lys Leu Asp Met Pro Tyr Glu Tyr Leu Arg Pro Ser Val Val

645 650 655

Glu Asn Ser Val Ala Ile

660

25 (2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Glu Leu Gln Xaa Trp Trp Tyr  
1 5

35 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Asp Val Trp Leu Leu Ala Lys  
1 5

5 (2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Gly Gln Phe Asp Ser  
1 5

15 (2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCAAGCCGAS SARTCRAAYT GNCC 24

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
25 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GAGCTTTGTC TCTGAAATAG TCAG 24

30 (2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTGAGGAATC AATAGCTTGA AGAG 24

(2) INFORMATION FOR SEQ ID NO:32:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GATGTGTGAC AGCCTCATGG ATG 23

(2) INFORMATION FOR SEQ ID NO:33:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CAAGCTTAGG AGGATGGCGA AATGCAGG 28

20 (2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGAATTCTATG TTAGATGGAG ACACGTGTT 28

(2) INFORMATION FOR SEQ ID NO:35:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Gly Gln. Tyr Asp Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:36:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Xaa Val Asp Trp Leu Leu Ala Ala Lys Xaa Trp Val Arg  
1 5 10

15 It will be understood that various details of the invention may be changed without departing from the scope of the invention. Furthermore, the foregoing description is for the purpose of illustration only, and not for the purpose of limitation--the invention being defined by the claims.